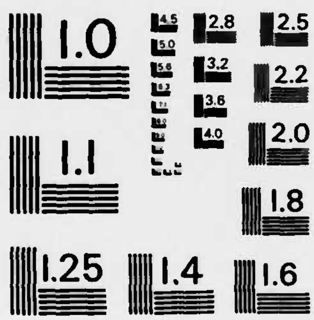
AD-A148 438 GENERATING POSITIVELY CORRELATED RANDOM VARIABLES FROM A SEQUENCE OF INDE..(U) FLORIDA STATE UNIV TALLAHASSEE DEPT OF STATISTICS M ABDEL-HAMEED ET AL. AUG 84 1/1

UNCLASSIFIED FSU-M685 AFOSR-TR-84-1092 F49620-82-K-0007 NL

END

FILMED

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

AD-A148 438

(4)

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S): FSU-M685	5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR- 84-1092
6a. NAME OF PERFORMING ORGANIZATION: Florida State University; 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: Air Force Office of Scientific Research
6c. ADDRESS (City, State and ZIP Code): Department of Statistics, Tallahassee FL 32306	7b. ADDRESS (City, State and ZIP Code): Directorate of Mathematical & Information Sciences, Bolling AFB DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: AFOSR; 8b. OFFICE SYMBOL (If applicable): NM	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: F49620-82-K-0007
8c. ADDRESS (City, State and ZIP Code): Bolling AFB DC 20332-6448	10. SOURCE OF FUNDING NOS.: PROGRAM ELEMENT NO. 61102F; PROJECT NO. 2304; TASK NO. A5; WORK UNIT NO

11. TITLE (Include Security Classification): GENERATING POSITIVELY CORRELATED RANDOM VARIABLES FROM A SEQUENCE OF INDEPENDENT RANDOM /VARIABLES WITH SYMMETRIC LOGARITHMICALLY CONCAVE DENSITIES

12. PERSONAL AUTHOR(S): Mohamed Abdel-Hameed and Frank Proschan

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Technical	FROM ____ TO ____	AUG 84	9

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Let $\underline{X} = (X_1,...,X_n)$ be independent random variables with logarithmically concave symmetric densities. The authors show that for any logarithmically concave functions f and g on R^n that are invariant under sign changes, $\text{Cov}(f(\underline{X}),g(\underline{X})) \geq 0$. Bounds on the values of logarithmically concave densities on R^n evaluated at the mean vector are also given.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL: MAJ Brian W. Woodruff	22b. TELEPHONE NUMBER (Include Area Code): (202) 767- 5027; 22c. OFFICE SYMBOL: NM

DD FORM 1473, 83 APR — EDITION OF 1 JAN 73 IS OBSOLETE.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

AFOSR-TR- 84-1092

GENERATING POSITIVELY CORRELATED RANDOM VARIABLES FROM A SEQUENCE OF INDEPENDENT RANDOM VARIABLES WITH SYMMETRIC LOGARITHMICALLY CONCAVE DENSITIES

by

Mohamed Abdel-Hameed[1]
University of North Carolina - Charlotte
and Kuwait University

Frank Proschan[2]
Florida State University
Tallahassee, Florida

Accession For	
NTIS GRA&I	X
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

UNCC Report No. 21

August, 1984

1. Research supported by the United States Air Force office of Scientific Research Under Grant AFOSR 80 - 0245

2. Research sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant (or cooperative agreement) Number AFOSR-82-K-0007. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon.

GENERATING POSITIVELY CORRELATED RANDOM VARIABLES FROM A SEQUENCE OF INDEPENDENT RANDOM VARIABLES WITH SYMMETRIC LOGARITHMICALLY CONCAVE DENSITIES

by

Mohamed Abdel-Hameed and Frank Proschan

ABSTRACT

Let $\underline{X} = (\underline{X}_1,\ldots,X_n)$ be independent random variables with logarithmically concave symmetric densities. We show that for any logarithmically concave functions f and g on R^n that are invariant under sign changes,

$$\mathrm{Cov}(f(\underline{X}),g(\underline{X})) \geq 0.$$

Bounds on the values of logarithmically concave densities on R^n evaluated at the mean vector are also given.

84 12 03 24 4

I. INTRODUCTION AND SUMMARY.

The area of positive and negative dependence of multivariate distributions has attracted the attention of many authors over the past decade. (See Abdel-Hameed and Sampson (1978), Jogdeo (1977), Kanter (1975), Karlin and Rinott (1980), Dykstra (1980), and Prekopa (1973).) Pitt (1977) proves that if n(x) is the standard normal density on R^2 and if A and B are balanced convex subsets of R^2 (i.e., $A=-A$ and $B=-B$) then

$$\int_{A\cap B} n(x)dx \geq \left(\int_A n(x)dx\right)\left(\int_B n(x)dx\right).$$

The question whether Pitt's result is true for any standard normal density on R^n, $n > 2$, remains unanswered.

In this paper we investigate covariance inequalities for a class of logarithmically concave densities. We show that if $X_1, \ldots, X_n$ are independent random variables each having a symmetric logarithmically concave symmetric density, then the random variables $Y_1 = f(X_1,\ldots,X_n)$ and $Y_2 = g(X_1,\ldots,X_n)$ are positively correlated whenever f and g belong to a certain class of logarithmically concave functions on R^n. In particular it follows that if A and B are subsets of R^n that are symmetric along all the axes, then

$$\int_{A\cap B} h(x)dx \geq \left(\int_A h(x)dx\right)\left(\int_B h(x)dx\right),$$

where h is the joint density of the independent random variables $X_1, \ldots, X_n$. We remark that a subset of R^n that is symmetric along all the axes is necessarily a balanced set and it follows that if

$(X_1,\ldots,X_n)$ is a standard normal vector, with density n then

$$\int_{A\cap B} n(x)dx \ge \left(\int_A n(x)dx\right)\left(\int_B n(x)dx\right)$$

for all subsets A and B of R^n that are symmetric along all axes.

Throughout, the word "symmetric" will be used to mean "symmetric about the origin". For n=1, 2, ..., let

$H_n = \{f: R^n \to R_+$, f is logarithmically concave and symmetric$\}$,

$A_n = \{A$: A is a $n\times n$ diagonal matrix with diagonal elements $\pm 1\}$,

$G_n = \{f\in H_n;\ f(\underline{x}A) = f(\underline{x})$ for all $\underline{x}$ in R_n and all $A\in A_n\}$, and

$L_n = \{K$: K is a convex symmetric subset of $R^n\}$.

<u>Section 2</u>. Positive Correlations of Functions of Multivariate Random Variables with Logarithmically Concave Densities.

In this section we will show that if X_1, ..., X_n are independent random variables each having a logarithmically concave symmetric density, then the random variables $Y_1 = f(X_1,\ldots,X_n)$ and $Y_2 = g(X_1,\ldots,X_n)$ are positively correlated whenever f and g belong to G_n.

<u>2.1 Theorem</u>. Let H be a convex subset of R^n. Then $f: H\to R_+$ is in H_n if and only if the set $H^+ = \{(\underline{x},z): f(\underline{x}) \ge e^z\}$ is a convex subset of R^{n+1} and $\{\underline{x} : f(\underline{x}) \ge a\}$ is a symmetric subset of R^n for each $a\in R_+$.

<u>Proof</u>. (If) Suppose that f is not in H_n. Then f is either not logarithmically concave or not symmetric. First assume that f is not logarithmically concave on H. Then there exists $\underline{x}_1$, $\underline{x}_2$ in H and a in (0,1) such that

$$f(a\underline{x}_1 + (1-a)\underline{x}_2) < f^a(\underline{x}_1)\, f^{1-a}(\underline{x}_2).$$

Thus the point $(a\underline{x}_1+(1-a)\underline{x}_2,\ a\,\ell n\, f(\underline{x}_1)+(1-a)\ell n\, f(\underline{x}_2)$ belongs to the line segment joining $(\underline{x}_1, \ell n\, f(\underline{x}_1), (\underline{x}_2, \ell n\, f(\underline{x}_2))$ but not in H^+. Since $(\underline{x}_1, \ell n\, f(\underline{x}_1))$ and $(\underline{x}_2, \ell n\, f(\underline{x}_2))$ are in H^+, then H^+ is not convex.

If f is not symmetric on H, then there exists $\underline{x}_0$ in H such that $f(\underline{x}_0) \neq f(-\underline{x}_0)$. Let $a_0=(f(\underline{x}_0) \vee f(-\underline{x}_0))$ and $K_{a_0}=\{\underline{x}:f(\underline{x}) \geq a_0\}$. Then either $\underline{x}_0$ or $-\underline{x}_0$ is in K_{a_0} but not both, contradicting the assumption that K_{a_0} is symmetric.

(Only if). Let $f \in H_n$, $(\underline{x}_1, z_1)$, $(\underline{x}_2, z_2)$ be any two points in H^+ and ℓ is the line joining them. Let $(\underline{x}, z_2)$ be any two points in H^+ and ℓ is the line joining them. Let $(\underline{x}, z)$ be any point on ℓ. Then there exists $0 \leq a \leq 1$ such that

$$\underline{x} = a\underline{x}_1 + (1-a)\underline{x}_2,$$
$$\underline{z} = az_1 + (1-a)z_2.$$

Since $f(\underline{x}_1) \geq e^{z_1}$ and $f(x_2) \geq e^{z_2}$, it follows that $f(\underline{x}) \geq e^{z}$. Therefore, $(\underline{x},z)$ is in H^+. Thus H^+ is convex.

The fact that $\{\underline{x}:f(\underline{x}) > a\}$ is symmetric for each $a \in R_+$ is immediate since f is symmetric.||

<u>2.2 Corollary</u>. Let H be a convex subset of R^n and assume that $f:H \to R_+$ is in H_n. Then, $\{\underline{x}:f(\underline{x}) \geq a\}$ is a convex and symmetric subset of R^n for each $a \in R_+$.

<u>Proof</u>: The symmetry of the set $\{\underline{x}:f(\underline{x}) \geq a\}$ is proved in Theorem 2.1 . Now assume that $\underline{x}_1$, $\underline{x}_2$ are in $\{\underline{x}:f(\underline{x}) \geq a\}$. Then for a in (0,1) we have

$$f(a\underline{x}_1 + (1-a)\underline{x}_2) \geq f^a(\underline{x}_1)\, f^{1-a}(\underline{x}_2) \geq a.$$

Thus, $a\underline{x}_1+(1-a)\underline{x}_2$ is in the set and hence it is convex; since a is arbitrary, the result follows.||

The converse to Corollary 2.2 is not true:

Let f be a concave symmetric function on R which is not in H_1. Then the sets $\{\underline{x}: f(\underline{x}) \geq a\}$ are convex symmetric subsets of R^1. However f is not in H_1.

The following lemma is due to Hoeffding and is a restatement of Lemma 2 of Lehmann [1966].

2.3 Lemma. Let X and Y be extended-valued random variables. Then

$$\text{Cov}(X,Y) = \int_{R^2} \text{Cov}\{I_{X^{-1}[x,\infty]}, I_{Y^{-1}[y,\infty]}\}dxdy,$$ where the

equality is valid even if one side is infinite.

2.4 Lemma. For any $K \in L_{n-1}$, $h \in H_{n-1}$, and $f \in H_n$, the function $g: R \to R_+$ defined by

$$g(x) = \int_K f(x_1,\ldots,x_{n-1},x)\, h(x_1,\ldots,x_{n-1})\, dx_1\ldots dx_{n-1}$$

is in H_1.

Proof: The logconcavity of g follows from Theorem 6 of Prékopa [1973]. The symmetry of g follows from the symmetry of f, h, and k. Thus g is in H_1, as desired.||

2.5 Theorem. Let f and g be in H_1. Suppose that (Ω,F,P) is a probability space and X is an extended-value random variable defined on (Ω,F,P). Then for each f, g in H_1 we have

$$\text{Cov}(f(X),g(X)) \geq 0.$$

Proof: By Lemma 2.3 we have

$$\text{Cov}(f(X),g(X)) = \int_{R^2} \text{Cov}\{I_{[x,\infty]}\circ f(X), I_{[Y,\infty]}\circ g(X)\}dxdy$$

since

$$I_{[x,\infty]}\circ f(X) = I_{f^{-1}[x,\infty]}\circ X.$$

From Corollary 2.2 we know that there exists a constant $a > 0$ such that $f^{-1}[x,\infty] = [-a,a]$. Thus there exists an $a > 0$ such that

$$I_{[x,\infty]} \circ f(X) = I_{[-a,a]} \circ X$$

Similarly we conclude that there is a $b > 0$ such that

$$I_{[y,\infty]} \circ g(X) = I_{[-b,b]} \circ X.$$

Therefore,

$$\mathrm{Cov}(I_{[x,\infty]} \circ f(X), I_{[y,\infty]} \circ g(X)) \text{ equals}$$

$$P\{X \in \{-[a \wedge b],[a \wedge b]\}\} - P\{X \in [-a,a]\} \cdot P\{X \in [-b,b]\},$$

which is clearly nonnegative. From Lemma 2.3 it follows that $\mathrm{Cov}(f(X),\ g(X))$ must be nonnegative.||

__2.6 Theorem__. Let $X_1, \ldots, X_n$ be independent random variables each having a density that belongs to H_1. Then for all f and g in G_n we have $\mathrm{Cov}(f(\underline{X}),\ g(X)) \geq 0$.

__Proof__: We proceed by induction on n. By Theorem 2.5 the result is true for $n=1$. Now assume the result is true for some n_0. For f and g in H_{n_0+1} and $\underline{X} = (X_1,\ldots,X_{n_0+1})$, we write

$$\mathrm{Cov}(f(\underline{X},g(\underline{X})) = E[\mathrm{Cov}\{f(X),g(\underline{X})|X_{n_0+1}\}]$$
$$+ \mathrm{Cov}[Ef(\underline{X}|X_{n_0+1}),\ Eg(\underline{X}|X_{n_0+1})].$$

For a fixed x_{n_0+1}, the function $f_{x_{n_0+1}} : R^n \to R_+$ is defined for any $\underline{x} = (x_1,\ldots,x_{n_0})$ by $f_{x_{n_0+1}}(x_1,\ldots,x_{n_0}) = f(x_1,\ldots,x_{n_0+1})$. For $\underline{x}_1 = (x_{11},\ldots,x_{1n_0})$, $\underline{x}_2 = (x_{21},\ldots,x_{2n_0})$, we let $\underline{x}_1^* = (x_{11},\ldots,x_{1n_0}, x_{n_0+1})$ and $\underline{x}_2^* = (x_{21},\ldots,x_{2n_0},x_{n_0+1})$. Then for a in $(0,1)$,

$$f_{x_{n_0+1}}(a\underline{x}_1+(1-a)\underline{x}_2) = f(a\underline{x}_1^*+(1-a)\underline{x}_2^*) \geq f^a(\underline{x}_1^*)f^{1-a}(\underline{x}_2^*)$$
$$= f^a_{x_{n_0+1}}(\underline{x}_1)\ f^{1-a}_{x_{n_0+1}}(\underline{x}_2)$$

Moreover, for any matrix A in A_{n_0} and $\underline{x} = (x_1,\ldots,x_{n_0})$ in R^{n_0} and $\underline{x}^* = (x_1,\ldots,x_{n_0},x_{n_0+1})$, we have $f_{x_{n_0+1}}(\underline{x}A) = f(\underline{x}^*A^*)$, where A^* = the $(n_0+1) \times (n_0+1)$ diagonal matrix defined by $A^* = \binom{A0}{01}$.

Since f is in G_{n_0+1}, then we have that

$$\begin{aligned} f_{x_{n_0+1}}(\underline{x}A) &= f(\underline{x}^*A^*) \\ &= f(\underline{x}^*) \\ &= f_{x_{n_0+1}}(\underline{x}). \end{aligned}$$

Therefore $f_{x_{n_0+1}}(x_1,\ldots,x_{n_0})$ is in G_{n_0}. The induction hypothesis combined with the above argument gives

$$\begin{aligned} &E[\mathrm{Cov}\{f(\underline{X}),g(\underline{X})\} \mid X_{n_0+1} = x_{n_0+1}] = \\ &E[\mathrm{Cov}\{f_{x_{n_0+1}}(\underline{X}^*),\ g_{x_{n_0+1}}(\underline{X}^*)\}] \geq 0 \end{aligned}$$

where $\underline{x}^* = (x_1,\ldots,x_{n_0})$ when $\underline{x} = (x_1,\ldots,x_{n_0+1})$.

From Lemma 2.4 and the hypothesis, we deduce that $Ef(\underline{X}|X_{n_0+1} = x_{n_0+1})$ as well as $Eg(\underline{X}|X_{n_0+1} = x_{n_0+1})$ are in H_1. By Theorem 2.5 we have $\mathrm{Cov}\{Ef(\underline{X}|X_{n_0+1}),\ Eg(\underline{X}|X_{n_0+1})\} \geq 0$. Thus we finally conclude that $\mathrm{Cov}(f(\underline{X}),g(\underline{X})) \geq 0$. ||

The proof of the following theorem can be obtained by imitating the proof of Theorem 2.6 .

2.7 Theorem. Let $X = (X_1,\ldots,X_n)$ be a standard normal vector with density n. Then

$$\int_{A\cap B} n(x)dx \geq \left(\int_A n(x)dx\right)\left(\int_B n(x)dx\right).$$

for all subsets A, B that are symmetric along all the axes.

Section 3. Bounds on Logarithmically Concave Densities.

In this section we derive some inequalities for strongly unimodal densities. First we define for $n=1,2,\ldots,$

$U_n = \{f:R^n \to R_+;$ f is a logarithmically concave density$\}$.

3.1 Lemma. Let f and g be functions mapping R^n into R_+. Then

$$\int f \ln(f/g) \geq (\int f) \ln(\int f/\int g).$$

In particular, if f is a density function on R^n, then

$$\int f \ln(f/g) \geq -\ln \int g,$$

for any measurable function $g:R^n \to R_+$.

Proof: First assume $\int f = \int g$. Then

$$\int f \ln(f/g) = -\int f \ln(g/f) \geq -\int f\{(g/f)-1\}$$

$$[\text{since } \ln x \leq (x-1),\ x \geq 0]$$

$$= \int f - \int g = 0.$$

Thus the inequality is satisfied in this case.

Now, if $\int f \neq \int g$, then define $g^* = (\int f/\int g)g$, and note that $\int f = \int g^*$. Therefore, using the above inequality we have $\int f \ln(f/g^*) \geq 0$. Using the definition of g^* and simplifying we get $\int f \ln(f/g) \geq (\int f) \ln(\int f/fg)$.||

3.2 Theorem. Let $\underline{X}$ be a random vector with density f belonging to U_n with finite mean $\underline{\mu}$. Let $g:R^n \to R$ be such that $\int \exp(-g) = 1$, then $f(\underline{\mu}) \geq \exp(-\int gf)$.

Proof: Take $f_1=f$, $f_2=e^{-g}$. By Lemma 3.1, we have $\int f_1 \ln(\frac{f_1}{f_2}) \geq 0$. Take $f_1=f$, $f_2=e^{-g}$. Using Jensen's inequality and the fact that f belongs to U_n, we deduce that

$$0 \le \int f\{\ell n\ f + g\} \le \ell n\ f(\mu) + \int fg,$$

completing the proof. ||

<u>3.3 Theorem</u>. Let $\underline{X}$ be a nonnegative random vector with density f belonging to U_n with finite mean $\underline{\mu}$. Then $f(\underline{\mu}) \ge \frac{1}{\mu_1 x \ldots x \mu_n} e^{-n}$. Equality is attained for $f(\underline{x}) = \prod_{i=1}^{n} \frac{1}{\mu_i} e^{-x_i/\mu_i}$ and therefore the bound is sharp.

<u>Proof</u>: Choose $g(x) = \sum_i \log \frac{\mu_i}{a_i} + \sum_i a_i \frac{x_i}{\mu_i}$ for $\underline{x} \ge 0$. Then $\int e^{-g} = \prod_{i=1}^{n} \int_0^\infty \frac{a_i}{\mu_i} e^{-a_i x_i/\mu_i} dx_i = 1$. Also $\int gf = \Sigma \log \frac{\mu_i}{a_i} + \Sigma a_i$. Thus by Theorem 3.2, $f(\underline{\mu}) \ge \prod_{i=1}^{n} (\frac{a_i}{\mu_i} e^{-a_i})$. The right hand side is maximized by choosing $a_i = 1$, $i = 1, \ldots, n$. Equality is attained for

$f(\underline{x}) = \prod_{i=1}^{n} \frac{1}{\mu_i} e^{-x_i/\mu_i}$, as may be directly verified. ||

The following theorem gives a lower bound on the peak of density function belonging to U_n in terms of the determinant of its covariance matrix.

<u>3.4 Theorem</u>. Let $\underline{X}$ be a multivariate vector with mean $\underline{\mu}$, covariance matrix Σ, and density f belonging to U_n. Then

$$f(\mu) \ge (2\pi)^{-n/2} |\Sigma|^{-1/2} \exp(-n/2).$$

<u>Proof</u>: Let $g(\underline{x}) = \frac{1}{2} \ell n |\Sigma| + (n/2) \ell n (2\pi) + \frac{1}{2}(\underline{x}-\underline{\mu})' \Sigma^{-1} (\underline{x}-\underline{\mu})$.

Then $\int e^{-g} = 1$, and $\int g(\underline{x}) f(\underline{x}) d\underline{x} = Eg(X) = \frac{1}{2}\ell n |\Sigma| + (n/2) \ell n (2\pi) + n/2$. The desired conclusion then follows from Theorem 3.2 . ||

REFERENCES

1. Abdel-Hameed, M.S. and Sampson, A. (1978). Positive dependence of the bivariate and trivariate absolute normal t, x^2 and F distributions. Ann. Statist., 6, pp. 1360-1368.

2. Joag-dev, K. (1977). Association and probability inequalities. Ann. Statist., 5, pp. 495-504.

3. Kanter, M. (1975). Stable densities under change of scale and total variation inequalities. Ann. Prob., 3, pp. 697-707.

4. Karlin, S. (1968). Total Positivity. Volume I. Stanford University Press.

5. Kowalczyk, T. and Pleszezyńska, E. (1977). Monotonic dependence functions of bivariate distributions. Ann. Statist., 5, pp. 121-1277.

6. Lehmann, E.L. (1966). Some concepts of dependence. Ann. Math. Statist., 37, pp. 1137-1153.

7. Pitt, L. (1977). A Gaussian correlation inequality for symmetric convex sets. Ann. Prob., 5, pp. 470-474.

8. Prékopa, A. (1971). Logarithmic concave measures with applications to stochastic programming. Acta Sci. Math, 32, pp. 301-316.

9. Prékopa, A. (1973). On logarithmic concave measures and functions. Acta Sci. Math., 34, pp. 335-343.

10. Van Zwet, W.R. (1964). Convex Transformations of Random Variables. Amsterdam: Mathematich Centrum.

END

FILMED

1-85

DTIC

DTIC